Research Paper | Computer Science | E-ISSN No : 2454-9916 | Volume : 3 | Issue : 5 | May 2017

CLASSIFICATION OF DATA USING MULTI-LEVEL AUTHENTICATION IN CLOUD COMPUTING

Amanpreet Singh [1] | Dr. Manju Bala [2] | Supreet Kaur [3]

[1] Professor, Department of CSE, ST. Soldier Institute of Engg. & Tech, Jalandhar.
[2] Professor, HOD of CSE Dept, Khalsa College, Amritsar.
[3] Assistant Professor, Lyallpur Khalsa College, Jalandhar.

ABSTRACT

Cloud computing is a new class of network based computing that provides the customers with computing resources as a service over a network on their demand. Cloud computing offers scalability, availability and different services as important benefits. Privacy-preserving information in cloud permits an information owner to send its encrypted information to a cloud server. Hence, security and privacy of knowledge is that the major concern within the cloud computing. To overcome this, various security aspects has been analyzed and then a framework to mitigate security issues at the level authentication and storage level in cloud computing is proposed. Encryption is a security technique which is widely used for data security. Also, a data classification approach based on data confidentiality is proposed. and implemented. Confidentiality, Integrity and Availability are the three main concepts which are taken into consideration while proposing the work. So, basically the main aim of this research work is to enhance the authentication security, to classify the data using machine learning algorithm, To Increase the confidentiality and security of the cloud computing and at last comparing the proposed ensemble learning scheme with KNN on the basis of Classification Accuracy, Precision, Recall, Data Encryption time, Data Decryption time.

KEYWORDS: Cloud Computing; security issues, privacy preserving, Integrity, confidentiality, availability, graphical passwords, Data classification, Machine Learning.

1. INTRODUCTION

This is the era of cloud computing. Cloud computing has dramatically shifted the technological advancements like the Internet have done 20-30years ago. Since then, the IT infrastructure has changed from desktop computers with 30MB to terabytes of online data storage. Due to the high availability and increased data connectivity, it is possible for users to access broad range of resources, services, software and applications with a click of mouse. Cloud computing is the internet based computing that provides "IT resources as a service" on demand of the user following the "pay-per-use". It employs parallel processing, distributed processing, grid computing and distributed database to enhance processing, in virtualization technology and the Internet broadband technology and based network [1].

Advancement of Cloud Computing is huge as for individual utilizations and employments related to business. Clients of distributed computing could use or keep up the online resources. Scalability, Virtualization, Mobility, Low Infrastructure Costs, and Increased Storage includes some of its benefits. So, Basically, Cloud Computing is recognized as a hottest technology which has a significant impact on IT field in the nearby future. In today's time Cloud computing is a fast growing field in computational research and industry. Cloud computing is the internet based computing that provides "IT resources as a service" on demand of the user following the "pay-per-use". The best examples of cloud service are Google App Engine, Gmail, Google Docs, Microsoft Windows Azure, and Amazon Elastic Compute Cloud (EC2). Cloud computing is a widespread term. As we compare it to a desktop computing, you may touch it, feel it. But the cloud is a real thing. It's a sort of imaginary, but yet it has great benefits. The cloud has three sections hardware, operating system and software. They all are basically rented over the Internet, in order to provide facilities to the clients. Cloud computing never focuses on the hardware structure rather it does emphasize over the code. Cloud computing automates delivery of selected cloud services to the users. It helps the organizations and individuals deploy IT resources at reduced total cost of ownership with faster provisioning.

To keep data secure in the cloud has always been a hefty issue in the IT's. But with technology that takes information outside the virtual secure walls, most corporations raised red flags. The usage of thin clients can actually be high jacked if people care less about the data. Because of the security concerns very few of the organizations are unwilling to present the case studies that are currently using the services. There are very few companies that are using the cloud computing at a very large scale. This leaves the organizations shy about the usage of cloud computing recourses even thought it has gained such a huge market. This paper concentrates on privacy issue in cloud computing. At whatever point the information is exchanged to the cloud server it experiences a security system i.e. encryption without comprehension the level of sensitivity of the data or the data is essentially put away on cloud server without securing it. [2] .To direct the security requirements of data, we have proposed a data classification model to classify the data according to its sensitivity level and then encrypting the only data which is required to secure using an encryption technique in cloud environment. Classification of objects is an indispensable field of research and of practical applications in numerous fields like pattern recognition and artificial intelligence, statistics, vision analysis and medicine. A very intelligent technique to secure the data would be to first classify the data into sensitive and non-sensitive data and then secure the sensitive data only. This will help to reduce the overhead in encrypting the entire data which will be exceptionally costly in connection of both time and memory. For encrypting the data many encryption techniques can be used and for classifying the data numerous classification algorithms are available in the field of data mining.

II. RELATED WORK

Shigeako, Manami, Motoi, Kanai et.al [3] have described the various risks a company faces while using cloud computing. The principle components from the social perspective, presence of at least two partners, security ensure in exposure condition and mission basic information issue. Chance framework strategy has been utilized to order hazards into four sorts i.e. hazard shirking, chance moderation, chance acknowledgment and hazard transference. Additionally a hazard counter measure table has been portrayed. On the premise of these countermeasures an order has been finished. ICT systems were analyzed for adaptable frameworks design, structures operation cost markdown, ecological effect diminish, and so on Cloud figuring has pulled in diversion as period that fathoms the ones. In the U.S., business undertaking Cloud offerings, comprising of Amazon EC2/S3, Google Apps, Force.Com, and Windows Azure, are picking up on expanding scope of clients. In addition, research of Cloud processing, alongside an administrative i-Japan approach and a begin of the smart Cloud take a gander at establishment of the Ministry of Internal Affairs and Communications, is advancing hurriedly in Japan. Solidly, the hazard issue from a man's perspective in such Cloud figuring is thoroughly removed with the hazard breakdown structure (RBS) technique. Besides, the danger components which have been removed are broke down and assessed. A point by point countermeasure and idea are created on the premise of these impacts. These in flip may be utilized to offer open Cloud utilize, give a lift to intensity by utilizing esteem deal, and development the effectiveness of organization manage.

Bokefode Jayant Det et.al [4] has instructed Cloud registering is one of the rising and promising region in Information Technology. It offers offerings to an organization over a group with the capacity to scale up or down their supplier necessities. Distributed computing offerings are set up and outfitted through an outsider, who having the framework. Distributed computing having scope of advantages however the most offices are stressed for tolerating it as a result of security issues and requesting circumstances having with cloud. Security necessities required at the association level powers to arrangement models that tackles the hierarchical and disseminated variables of measurements utilization. Such models need to give the security suggestions intended to watch certainties contrary to unapproved get admission to and change put away in a cloud. The proposed works of art depicts the method for displaying the assurance prerequisites from the state of mind of framework elements and commitments done in a venture by method for the utilization of making utilization of the cryptography

thoughts to store information on cloud with the littlest measure of time and expense for encryption and decoding methodologies. In this depictions, we utilized RSA and AES calculation for encryption and decoding of insights and capacity based thoroughly motivate admission to control model is utilized to offer get right of section to in accordance with the position played through client. This paper moreover shows the numerical model for figuring the consider of the client. This model gives the transferring rights to the benefactor while he/she pushed by means of way of the Administrator and Owner in the meantime as customers surpasses the predetermined involvement and takes conveyance of as genuine with limit cost.

Xuefeng Liu, Yuqing Zhang et.al [5] has introduced a very brilliant idea of multi-owner data sharing technique in the cloud. The use of encryption techniques, making the system more secure from the culprits has been illustrated. They implied that any user in the group can securely share the data with others. Explanations of Plutus and group signatures have been beautifully explained. The proposed system Mona was brilliant until and unless a flaw was detected. The missing of a backup plan is a point that can lead to a new enhanced system. The system description has been elaborated but seems a bit difficult to understand. The technique of sharing data with others in the group seems very eye catching.

Frank Simorjay et.al [6] has proposed the classification of organizing the data according to their sensitivity and how should the data be stored. Data classification has been used since decades in large organizations such as Microsoft, governments and military basis. Data classification fundaments authentication and authorization with the roles in cloud are interesting to read. The classification process is one of the effective ways to implement data classification. Ways to protect confidential data are also illustrated. A brief hint of data protection techniques is provided that can be a roadmap of changes.

Ming-Huang Guo et al. [7] has worked on graphical passwords for authentication system. In this they have used, two methods which are cloud devices and cloud environment. Here, main purpose is to authenticate users from authentication server. Here, the user using their cloud device first run cloud and then enter their user ID and the program shows has some graphs from where user selects a point on the graph. Now program generates public and private key which applies when the message is transmitted. Now, when the message is generated by user onto cloud new time stamp is added in the transmission. Now each time when message is retrieved from the cloud, the hash values are used to validate the message.

Liu Hao et.al [8] discovered a system in which they design cloud security storage system (CSSS). They also designed an interface by which user can safely access to the database. It provides the proper security features to the customer. They introduced the idea of distributed computing and other related information to raise the point information security. That implies how to securely spare information in the cloud. They directed inside and out exploration on the security issue of cloud information stockpiling, and outlines a Prototype framework for private distributed storage, which incorporated private cloud information insurance including access control, I/O characterization, metadata assurance for delicate information, and infection identification, content filtration, continuous reinforcement and quick data recovery, giving establishment to reviewing information stockpiling and insurance in private distributed storage framework for big business. Created secure-distributed storage framework is in trial, enhance and impeccable stage, and need experience the procedure of continually moving forward. Next, after further trial utilizes as a part of numerous associations and change the bug, they continuously acquainted their items with more schools and associations. In this way, this task should make great social and financial advantages. As the framework utilizes the substance based capacity techniques, intranet clients can get to required documents much quicker, and download records all the more advantageously, and improve their learning activity on the Cloud.

III. PROPOSED METHODOLOGY

The paper proposes a secure data classification model using novel boosting supervised machine learning approach. In this, data is classified according to its sensitivity level. Then encrypting only, the data which is required to be secure using a hybrid privacy preserving based image steganography technique in cloud environment. [9] The proposed work also ensures the privacy and integrity of data using hashing approach [10].

Step 1: Authentication Level:

a) ***Owner:*** top level security like giving the access after finger print scan, various security questions.

b) ***Administrator:*** second level security i.e. after asking various security questions then provides the access.

c) ***User:*** third level security i.e. providing access after username and graphical password matching.

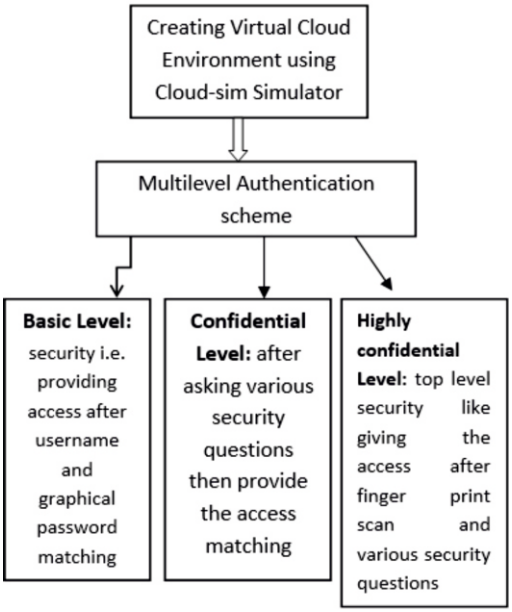

Figure 1: Authentication Level

Step 2: Data Classification

The algorithm is divided into 2 phases. In the first phase, each of the base level classifiers takes part in the j- fold cross validation training where a vector is returned in the form <(y'0… y'm), yj> where y'm is the predicted output of the mth classifier and yj is the expected output for the same. In the second phase this input is given for the Meta learning algorithm which adjusts the errors in such a way that the classification of the combined model is optimized.

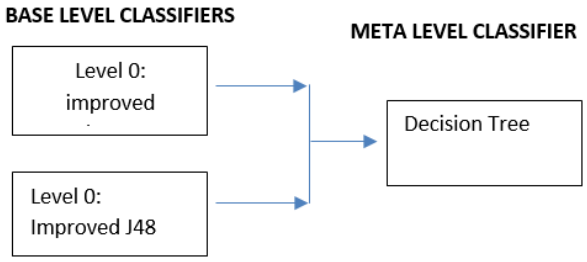

Figure 2: Ensemble learning basic model

Ensemble Learning Procedure

Algorithm 1: *Pseudo-Code of Main Function*

Procedure 1: Selecting base and Meta classification layers

Input: original-dataset: Dataset, folds: Integer

{

Dataset ←original-dataset

define Array-Of-Classifiers as array of classifiers that contains Improved Adaboost and Improved J48

For layer = 0 to 2 do:

{

For each fold in folds do:

{

If layer ≠2 do:

Array-Of-Classifiers ← Train_SingleLayer (layer, train-set, folds) // Here we are calling classification algos

New-Instances ← Classify (test-set, Array-Of-Classifiers)

Add New-Instances to dataset [layer+1]

Else

Train_SingleLayer (layer, train-set, folds)

}

Layer = layer + 1

}

Train-Single-Layer (layer0, original-dataset) //Rebuild Base classifiers using // the original dataset

}

Algorithm 2: *Pseudo code of training process for each one of its base level layer, which is systematically called by the Main Procedure*

Procedure 2: Classifying a Single layer

Train-Single-Layer Input: Layer

Number: Integer, dataset: Dataset, folds: Integer

Output: Successor-Dataset: Dataset

{

Successor Dataset ← empty Group

For each fold in folds do:

Build Classifiers (Layer Number, train-set)

For each instance in test-set

{

Produce probabilities-vector by applying instance on current layer's classifiers.

Generate a new Instance from probabilities-vector

Add the new Instance to Successor-Dataset

}

}

Return Successor-Dataset

}

Step3: Data hiding Architecture
Proposed algorithm for encoding the highly confidential data

Inputs: Dataset, RGB image

Output: Encrypted Image

Step1: Take the RGB image as input

Step2: Detect the edges by applying canny edge detection method.

Step3: The pixels constitute the edges are denoted by X, i, j constitute the dimensions of the pixel such that P (i, j, X). Put them into an array A1.

Step4: Read character from Dataset that is to be masked and convert the ASCII value of the character into equivalent binary value of 8 bits.

Step5: for each message bit m initialize random function generator and randomly select index value of array A1.

Step6: Extract the pixel value from that index and check its LSB; if LSB matched with the message bit then mask the bit otherwise ignore and find the index value again by random method.

Step7: store the randomly generated index values into a text file as a key and send the image to cloud.

RSA Algorithm:
The RSA algorithm involves four steps: key generation, key distribution, encryption and decryption.

Using an encryption key (e,n), the algorithm is as follows:

1. Represent the message as an integer between 0 and (n-1). Large messages can be broken up into a number of blocks. Each block would then be represented by an integer in the same range.

2. Encrypt the message by raising it to the eth power modulo n. The result is a cipher text message C.

3. To decrypt cipher text message C, raise it to another power d modulo n

The encryption key (e,n) is made public. The decryption key (d,n) is kept private by the user.

IV. PERFORMANCE PARAMETERS
The evaluation parameters considered for evaluating the performance of the proposed system are:

a. Classification Accuracy

b. Classification Time

c. Data Encryption time

V. RESULTS AND DISCUSSIONS

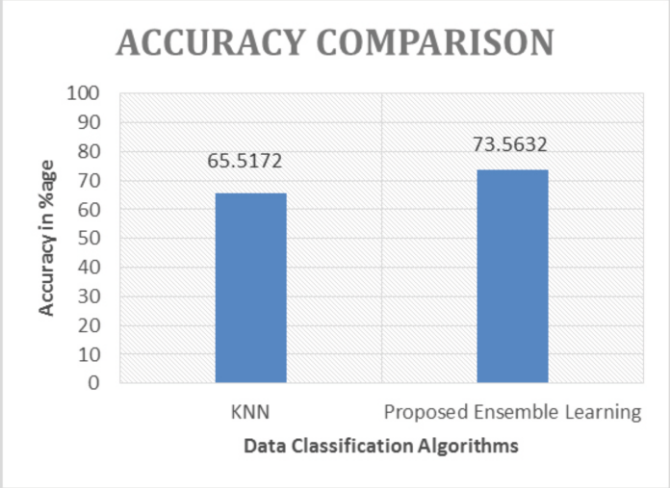

Figure 3: Performance analysis of data classification algorithms on the basis of accuracy

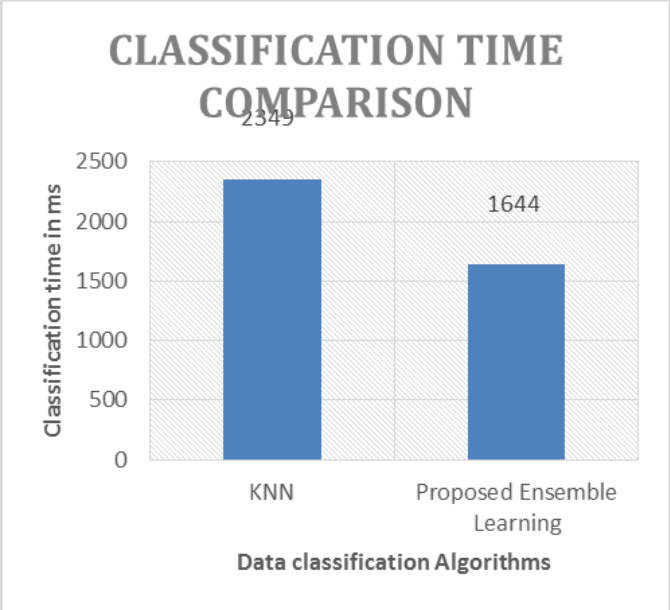

Figure 4: Classification Time Comparison of KNN Algorithm with the Proposed Ensemble Algorithm

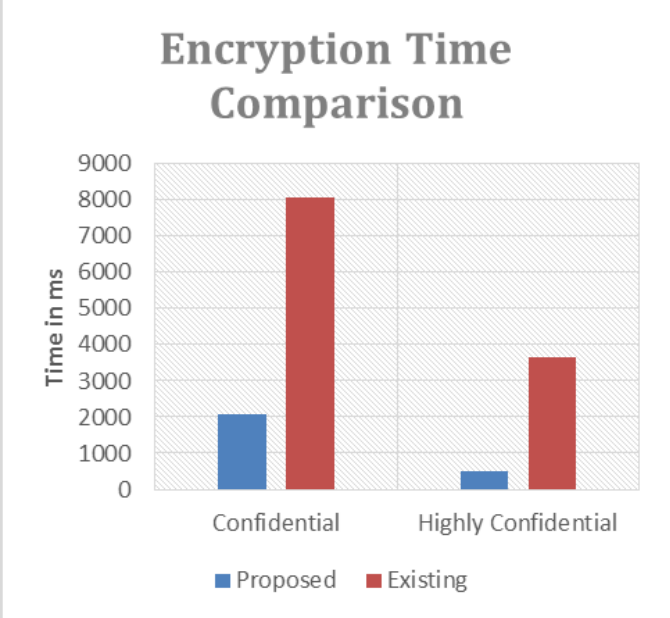

Figure 5: Encryption time Comparison of Existing techniques with the Proposed Technique on the basis of data classification labels.

The above figure shows the performance analysis of the proposed methodology with the previous method. It is clearly analyzed from the performance graphs that the proposed technique is better than the previous approach. If we talk about accuracy parameter in figure 3, KNN algorithm is having accuracy 65.5% and improved boosting is having 73.56% i.e. proposed algorithm has classified data more correctly. Similarly, in case of classification time in figure 4, proposed ensemble learning technique takes 1644 milliseconds and the ANN algorithm takes 2349 milliseconds to classify the data. Also, figure 5 is predicting comparison of encryption time on the basis of confidential and highly confidential. So, we can say that the proposed methodology performs betters in respect to all the above parameters.

VI.CONCLUSION

In this article, a technique for data confidentiality in cloud environment is proposed. The focus of the research was to characterize the data taking into account the security prerequisites of the information that divides the data into sensitive and non-sensitive using improved machine learning algorithm. The fundamental contribution of this security model is data confidentiality and classification of data using machine learning classification approach. Furthermore, to enhance the security at the authentication level, multilevel authentication scheme has been used based on the different types of users including owner, admin and user each one has different levels of security also in that image sequencing passwords based on different themes has been used in order to avoid un-authorized access to the cloud environment. Also the results show that the proposed ensemble learning technique works better than the K-NN classification technique in terms of both the accuracy and the classification time.

ACKNOWLEDGMENT

The paper has been composed with the kind assistance, guidance and support of my department who have helped me in this work. I would like to thank all the people whose encouragement and support has made the fulfillment of this work conceivable.

REFERENCES

[1] Stuti Srivastava,Prem Sewak Sudhish, "Security in cloud computing systems: A review of challenges and solutions for security in distributed computing environments,"2015 39th National Systems Conference (NSC), pp. 1 - 5, 2015.

[2] Abdullah, A., Hashim, F., & Al-Haddad, S., "A review of cloud security based on cryptographic mechanisms", IEEE, Kuala Lumpur, pp. 106-111,2014.

[3] Shigeaki Tanimoto, Manami Hiramoto, Motoi Iwashita, Hiroyuki Sato, Atsushi Kanai (2011), "Risk Management on the Security Problem in Cloud Computing", Chiba Institute of Technology, Japan; The University of Tokyo, Japan; Hosei University, Japan

[4] Bokefode Jayant D., UbaleSwapnaja A., Pingale Subhash V., Karande Kailash J. and ApateSulabha S... "Developing Secure Cloud Storage System by Applying AES and RSA Cryptography Algorithms with Role based Access Control Model", International Journal of Computer Applications 118(12):46-52, May 2015.

[5] Xuefeng Liu, Yuqing Zhang (2013), "Mona: Secure Multi-Owner Data Sharing for Dynamic Groups in the Cloud", National Key Laboratory of Integrated Services Networks, Xidian University, No. 2, Taibai Road, Xian city, China; National Computer Network Intrusion Protection Center, Graduate University of Chinese Academy of Sciences, No. 19, Yuquan road, Beijing, China

[6] Frank Simorjay (2014), "Data classification ffor cloud readiness", Microsoft Trustworthy Computing

[7] Gaurav, S. M., Khochare, N., &Rane, P. (2014) "Graphical Password Authentication". International Conference on signal processing systems, IEEE, Nagpur pp. 479-483.

[8] Liu Hao and Dezhi Han, "The study and design on secure-cloud storage system",International Conference on Electrical and Control Engineering (ICECE),pp. 5126-5129, 2011.

[9] Chen, D., & Zhao, H., "Data Security and Privacy Protection in cloud computing." IEEE, Hangzhou, pp. 647-651,2012.

[10] Mohammed Faez Al-Jaberi and Anazida Zainal, "Data Integrity and Privacy Model in Cloud Computing" International Symposium on Biometrics and Security Technologies, IEEE, pp.280-284, 2014